



Tema:

Machine Learning aplicado à análise de pênaltis

Introdução

De um modo geral este trabalho apresenta os resultados da aplicação de Machine Learning, em imagens, para a interpretação da previsibilidade de resultados de cobranças de pênaltis. São estudados também os parâmetros que são mais relevantes para a acertabilidade dos diferentes algoritmos estudados, bem como são feitas as interpretações sobre os resultados encontrados.

Cobranças de pênaltis

Um momento crucial em partidas de futebol é o momento da cobrança de um pênalti. O pênalti é o momento em que um atleta tem a chance de chutar para o gol sem a interferência de adversários, a não ser pelo goleiro.

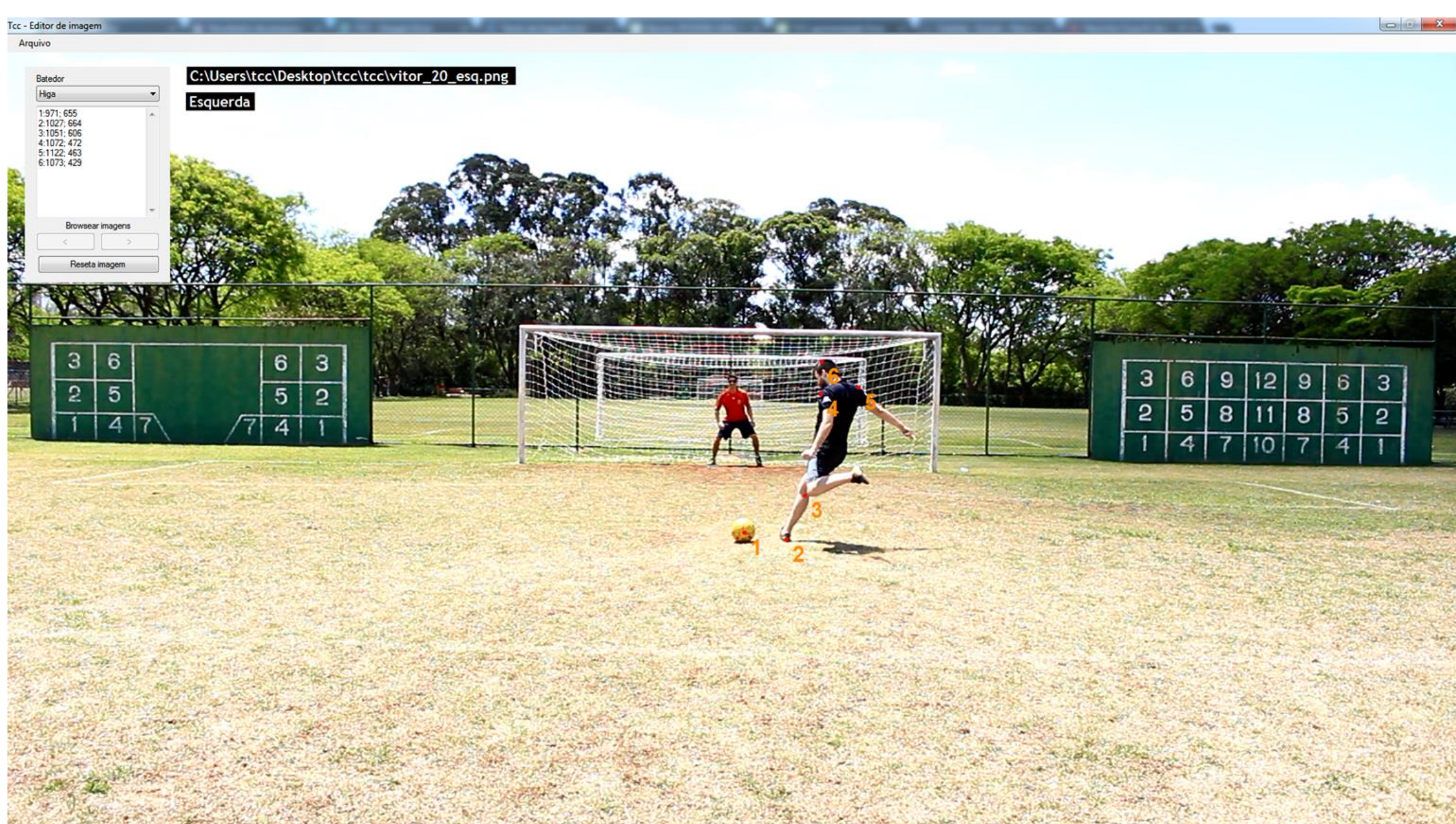
Uma cobrança de pênalti bem executada é praticamente infalível, dado a desvantagem do goleiro em relação ao bater, em termos de tempo de reação e tamanho da baliza.

Conduzimos, portanto, um estudo sobre a possibilidade de previsão, via Machine Learning, de qual será o lugar do gol em que o atacante irá chutar a bola, de forma a avaliar se certos bateres são mais previsíveis que outros, e quais os algoritmos mais adequados para esse fim.

Extração de parâmetros

Diferentes bateres tem diferentes formas de abordar uma cobrança de pênalti. Uma primeira tarefa, portanto, consistiu em entender quais seriam os pontos de referência de cada jogador que diriam, com maior probabilidade de acerto, qual seria o produto da cobrança: um chute para a direita ou para a esquerda.

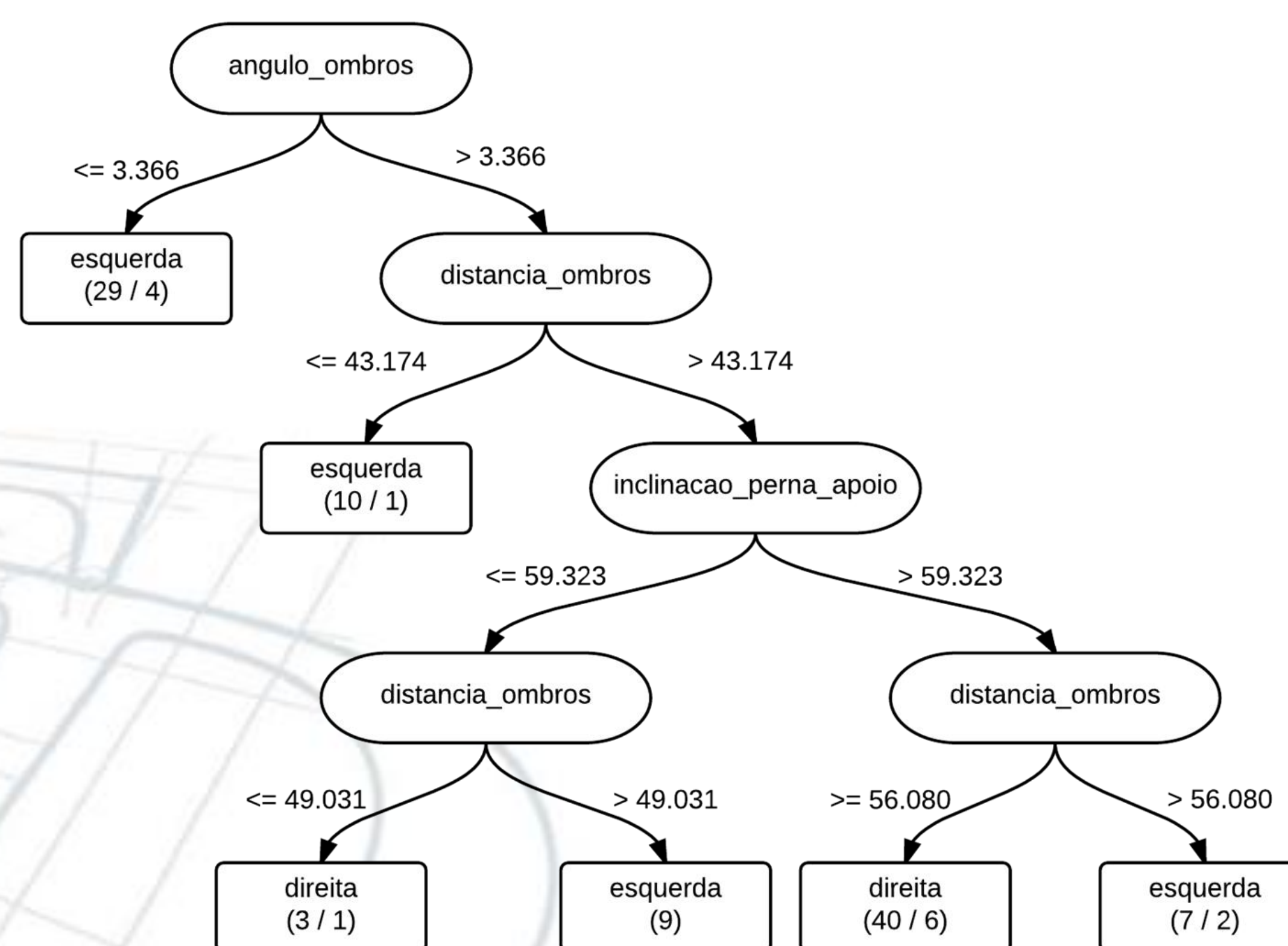
Um programa foi desenvolvido para automatizar a extração desses parâmetros para interpretação posterior, que consiste em compilar os dados da cobrança, como distância do pé de apoio do jogador à bola, inclinação de seu corpo, entre outros fatores, para posterior análise com ferramentas e algoritmos de Machine Learning.



Machine Learning

Com os parâmetros identificados e devidamente extraídos, foram feitas análises dos inputs em relação à diversos algoritmos diferentes de Machine Learning, para avaliação de qual seria mais adequado para estudo dos dados.

Árvore de decisão, Redes Neurais, Classificador Naive Bayes e Support Vector Machine e Clusterização K-Means foram os algoritmos estudados, cada um oferecendo diferentes resultados com diferentes precisões e coeficientes.



Árvore de decisão

Interpretação dos resultados

Depois de rodados os algoritmos, os resultados foram analisados com matrizes de confusão (tabela de contingências), análises de sensibilidade, e cálculo dos coeficientes de performance para cada caso, quando aplicável.

Com isso foi possível tirar diferentes conclusões da performance (Precisão, Revocação, Medida-F, Acurácia, Especificidade) de cada um dos algoritmos, bem como identificar possíveis otimizações e falhas de adaptação deles com o propósito ao qual foram submetidos.

	Classificado: Esquerda	Classificado: Direita
Chute: Esquerda	39	16
Chute: Direita	13	30

Matriz de confusão / Tabela de contingência para árvore de decisão

Classe	% V positivo	% F positivo	Precisão	Revocação	Medição-F	ROC Area	PRC Area
Esquerda	0,709	0,302	0,750	0,709	0,729	0,692	0,696
Direita	0,698	0,291	0,652	0,698	0,674	0,692	0,606
Média Ponderada	0,704	0,297	0,707	0,704	0,705	0,692	0,657

Análise de coeficiente de performance para árvore de decisão

Integrantes: Davi Alvares Fochi Silveira
Eduardo Fujihara Zatz
Vitor Mendes Julio

Professor Orientador: Andre Riyuiti Hirakawa